

An Experimental Study of Cohesion and Coupling Metrics

Chung-Hong Lung
Department of Systems & Computer Eng.
Carleton University, Ottawa, Ontario Canada
email: chlung@sce.carleton.ca

Nishith Goel and Pulak Dhar
Cistel Technology Inc.
Ottawa, Ontario Canada
email: ngoel@cistel.com and pulak@cistel.com

Qualities Related to Cohesion and Coupling

- Survivability
- Maintainability
- Verifiability
- Flexibility
- Portability
- Reusability
- Interoperability
- Expandability

Metrics & System Studied

Metrics:

- H. Dhama's C&C metrics with modifications
J. of Systems and Software 1995
- D. Card and R. Glass's Data Complexity
Measuring Software Design Quality, Prentice Hall, 1990
- Relative Complexity
- LOC (Lines of Code)
- Cyclomatic Complexity
- LSS (Logical Source Statement)
- Others: Average Depth, Nesting Level, etc.

System:

Large-scale telecommunications software system

Dhama's Cohesion & Coupling Metrics

- Functional Cohesion
- Data Flow Cohesion
- Action-Bundling Cohesion
- Logical Bundling Cohesion
- Coupling

C&C Metrics Used in the Study

- Functional Cohesion: slight modification
- Data Cohesion: hybrid of data flow & action-bundling cohesion
- Uses Cohesion: density of variables
- Coupling: slight modification
- Data Complexity
- Logical Bundling Cohesion: similar to logical nesting level, not used.

Functional Cohesion

$F = 1/p,$

where $p = i1 + q_1i2 + u1 + q_2u2 + l1 + q_3l2 + g1 + q_4g2 + q_5w$

$i1$ = in data parameters

$i2$ = in control parameters

$u1$ = out data parameters

$u2$ = out control parameters

$l1$ = number of local variables used as data

$l2$ = number of local variables used as control

$g1$ = number of global variables used as data

$g2$ = number of global variables used as control

w = number of modules called

$q_1, q_2, q_3, q_4,$ and $q_5 = 2$

A variable could be used as both data & control.

A parameter could be both in & out parameter.

Data Cohesion

- Interdependencies among the different statements depending on the processing of data.
- Hybrid of Dhama's Data-flow and Action-bundling cohesion.
- Consider the type of statement and position of the variable.
- Data-flow cohesion occurs between two statements if data used in one statement is transformed and then used by another transformation or data in another statement.
- Action-bundling cohesion occurs when several actions are performed on a single piece of data which results in that data being transformed.
- Data cohesion (hybrid of data-flow & action-bundling) measures the number of statements (distance) separating pairs of statements that have either cohesion.

Data Cohesion

Example:

Data flow cohesion: B is transformed and used in another statement.

B = A	B = A	B = A
...
C = B	if (B < D)	write (B)

Action-bundling cohesion: C is used in both statements on the right.

B = C	B = C	B = C
...
D = C	for i in 1..C	write (C)

Uses Cohesion

- Uses cohesion measures code density and involves the number of local variables and global variables divided by the number of tokens (variables + constants + function calls) in the code.

Coupling

$F = 1/p,$

where $p = i1 + q_1i2 + u1 + q_2u2 + g1 + q_4g2 + w + r$

$i1$ = in data parameters

$i2$ = in control parameters

$u1$ = out data parameters

$u2$ = out control parameters

$g1$ = number of global variables used as data

$g2$ = number of global variables used as control

w = number of modules called

r = number of modules calling the module under consideration

$q_1, q_2, q_3,$ and $q_4 = 2$

A variable could be used as both data & control.

A parameter could be both in & out parameter.

Data Complexity

- Data Complexity

$$D(i) = V(i) / f(i) + 1$$

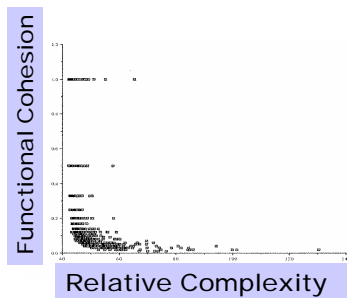
where

$D(i)$ = data complexity of module i

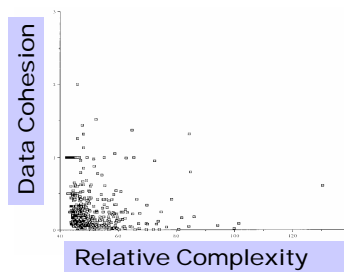
$V(i)$ = I/O variables in module i

$f(i)$ = fanout of module i

Functional Cohesion vs.
Relative Complexity

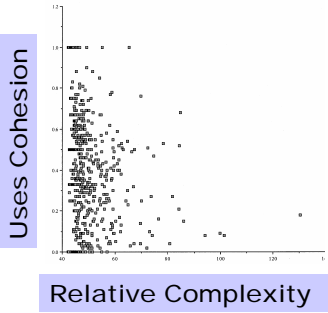


Data Cohesion vs.
Relative Complexity

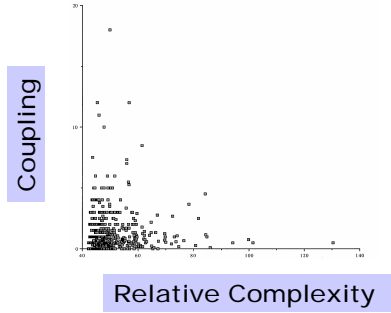


Each point represents a module which consists of one or more files

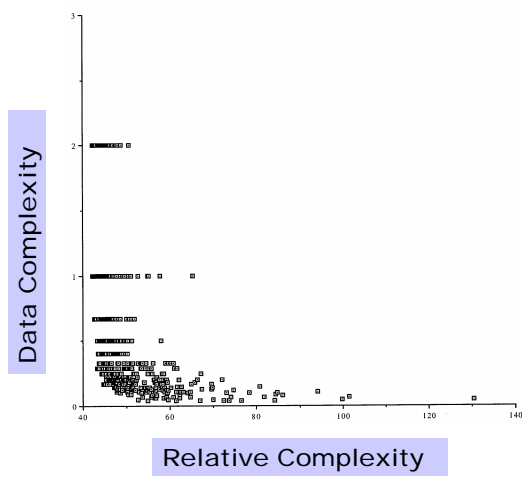
Uses Cohesion vs. Relative Complexity



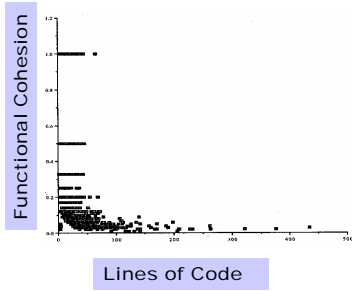
Coupling vs. Relative Complexity



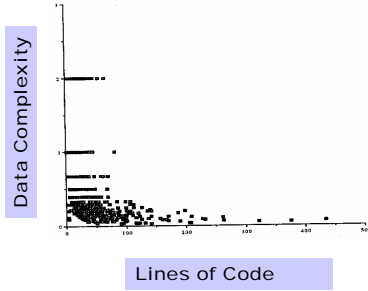
Data Complexity vs. Relative Complexity



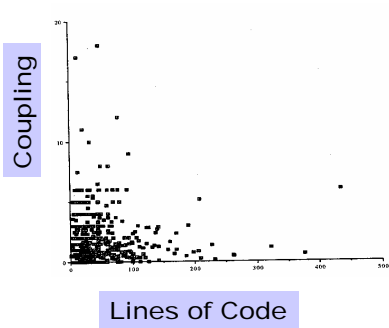
Functional Cohesion vs. LOC



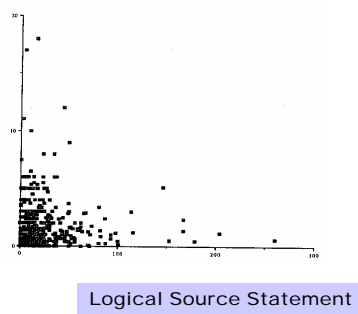
Data Complexity vs. LOC



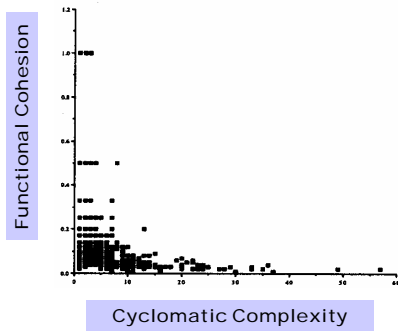
Coupling vs. LOC



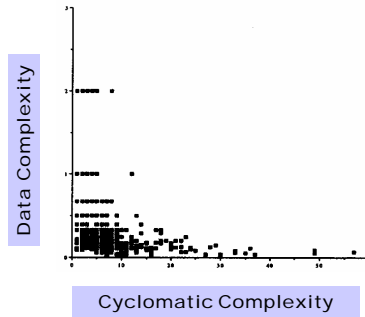
Coupling vs. LSS



Functional Cohesion vs.
Cyclomatic Complexity



Data Complexity vs.
Cyclomatic Complexity



Conclusions

- Functional cohesion, data complexity, and coupling seem to be consistent enough to produce detectable trends.
- LOC and Cyclomatic complexity seem correlate well with functional cohesion and data complexity.
- Other metrics studied that also correlate well with functional cohesion and data complexity include
 - Logical Source Statement
 - Physical Source Statement
 - Nesting Level
 - Average Depth